# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2002-073680

(43)Date of publication of application : 12.03.2002

(51)Int.Cl.
G06F 17/30

(21)Application number : 2000-261366

(22)Date of filing : 30.08.2000

(71)Applicant : MITSUBISHI RESEARCH INSTITUTE INC

(72)Inventor : NAKAMURA TATSUO

(54) TECHNICAL INFORMATION RETRIEVAL SYSTEM

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a technical information retrieval system which can retrieve the technical information in a simple procedure with high accuracy and with no skill required.
SOLUTION: This system includes a user terminal equipment 1 which inputs the retrieval statements and also outputs the retrieval results, a retrieval equipment 2 which retrieves the technical information from the retrieval statements and a management terminal equipment 3 which registers the technical information on the equipment 2. The morpheme analysis is applied to the retrieval statement inputted by a user for generating a word vector and the inner product is calculated between the word vector and the word vectors of each technical information registered in a technical information data base device 4. On the basis of these calculation results, the resemblance is decided among those word vectors so that the resemblance can be easily and highly accurately decided to each technical information. Furthermore, as the retrieval results are shown in a form where the degrees of resemblance can be understood, the degree of resemblance of each technical information can be objectively known despite the retrieval of plural pieces of technical information.
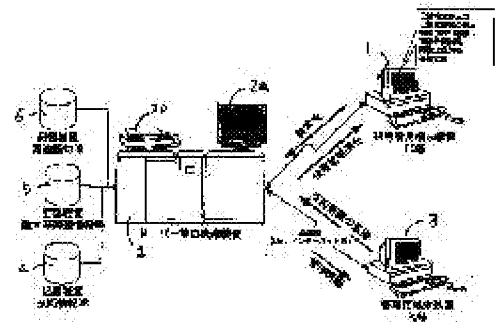
## DETAILED DESCRIPTION

[Detailed Description of the Invention]

[0001]

[Field of the Invention]This invention relates to the technical-information search system which searches the technical information based on the inputted retrieval sentence.

[0002]

[Description of the Prior Art]These days, a lot of technical information is electronized by making patent information into the start, and it is becoming accessible via the Internet etc. Compares with what is provided for nothing, specific ID, or a password, and some technical information is provided for pay, and the language and the data format of the technical information are also various. For this reason, the art of searching the desired technical information simply and promptly out of these technical information of a lot of is needed.

[0003]

[Problem(s) to be Solved by the Invention]When technical information, such as patent information, was searched, it searched by specifying a keyword conventionally in many cases. However, there is a problem that the accuracy of search does not go up unless it inputs an exact keyword.

[0004]For example, if it searches by specifying a general-purpose keyword, the applicable number of cases will increase rapidly and extraction of very required information will become difficult. Therefore, since only inaccurate search was completed even if existence of the skilled searcher is indispensable and ordinary persons searched, in order to have raised the accuracy of search, it had fallen into the vicious circle that the search system itself is seldom used.

[0005]On the other hand, the example which displays advertisement information on a homepage has increased with the spread of the Internet. When displaying the search results corresponding to the search condition which the user inputted as one of examples, the advertising grant system which displays the advertisement information which has relation in the search condition is proposed.

[0006]Beforehand, an advertiser chooses the classification of the product which is an advertised object, or service from the classifications on a system management, and registers it, and when this system displays search results, it displays simultaneously the advertisement information belonging to the classification of those search results.

[0007]However, in order to raise the goodness of fit of the product and service which are advertised objects, and a classification, the layered structure of a classification must be made finer. For this reason, the burden by the side of an advertiser will increase rapidly, and the advertising number itself will decrease.

[0008]When many products belonging to the same classification and information on service exist, a rational means to determine the priority of a display is not proposed.

[0009]An advertiser is thought of [ wanting to specify a wide range classification in many cases and ] so that as many advertisement information about a self product or service as possible may be displayed, but if it is made such, a user will be provided with a lot of advertisement information, and the merit which carries out classification of the advertisement information will fade.

[0010]It was in use that the license fee based on the contract concluded between the parties estimated conventionally as a technique by which the market value of technical information like a patent is evaluated. However, other than the party concerned, the contents of the licensing agreement were disclosure in many cases, and only evaluation of them about the existing dealings of the limited fields was completed, and they were not able to be applied as a forecasting method of the market value of new art.

[0011]this invention is made in view of such a point, and comes out. It is in providing the technical-information search system in which highly precise search is possible in a simple procedure, without requiring the purpose.

[0012]
[Means for Solving the Problem]This invention is provided with the following in order to solve SUBJECT mentioned above.

A related term extraction means to extract a related term out of an inputted retrieval sentence.

The 1st morphological-analysis means that conducts a morphological analysis to an extracted related term.

The 1st vector generating means that performs weighting which considered the contents of technical information to each word obtained by morphological analysis by said 1st morphological-analysis means, and generates vector information.

The 2nd morphological-analysis means that conducts a morphological analysis to two or more retrieval object technical information, As opposed to each word obtained by morphological analysis by said 2nd morphological-analysis means, The 2nd vector generating means that performs weighting which considered the contents of technical information and generates vector information, An inner-product-computing means to calculate an inner product between vector information generated by said 1st vector generating means, and vector information generated by said 2nd vector generating means, A similarity judging means which judges similarity of a retrieval sentence and retrieval object technical information based on the result of an operation of said inner-product-computing means, An output object selecting means which chooses an output object of a decision result by said similarity judging means, and a search-results output means which arranges a selected output object in order of the similarity judged by said similarity judging means, and outputs it.

[0013]Judging similarity of vector information corresponding to a related term extracted out of a retrieval sentence, and vector information of retrieval object technical information in this invention, arranging in order of similarity, and having made it output search results A sake, It can be grasped quantitatively whether it is a retrieval sentence and how much similar, and technical information similar to a retrieval sentence can be searched simply and promptly.

[0014]A related term extraction means to extract a related term out of a retrieval sentence into which this

invention was inputted, The 1st morphological-analysis means that conducts a morphological analysis to an extracted related term, As opposed to each word obtained by morphological analysis by said 1st morphological-analysis means, The 1st vector generating means that performs weighting which considered the contents of technical information and generates vector information, The 2nd morphological-analysis means that conducts a morphological analysis to two or more retrieval object technical information, As opposed to each word obtained by morphological analysis by said 2nd morphological-analysis means, The 2nd vector generating means that performs weighting which considered the contents of technical information and generates vector information, The 1st information accumulation means that stores vector information generated by said 2nd vector generating means, As opposed to each word obtained by morphological analysis by the 3rd morphological-analysis means that conducts a morphological analysis to advertisement information, and said 3rd morphological-analysis means, The 3rd vector generating means that performs weighting which considered the contents of technical information and generates vector information, and the 2nd information accumulation means that stores vector information generated by said 3rd vector generating means, An inner product is calculated between vector information stored in said 1st information accumulation means, and vector information generated by said 1st vector generating means, And an inner-product-computing means to calculate an inner product between vector information stored in said 2nd information accumulation means, and vector information generated by said 1st vector generating means, It has a similarity judging means which judges similarity based on the result of an operation of said inner-product-computing means, and a search-results output means which outputs advertisement information whose similarity is higher than the specified quantity while outputting retrieval object technical information in order of similarity based on judged similarity.

[0015]In this invention, since advertisement information was provided taking similarity of a retrieval sentence and retrieval object technical information into consideration, only those interested can be provided with advertisement information and advertising effectiveness can be heightened.

[0016]

[Embodiment of the Invention]Hereafter, the technical-information search system concerning this invention is explained concretely, referring to drawings.

[0017](A 1st embodiment) Drawing 1 is a block diagram showing the outline composition of a 1st embodiment of the technical-information search system concerning this invention. The technical-information search system of drawing 1 is provided with the following.

The terminal unit 1 for users which performs the input of a retrieval sentence, and the output of search results.

The retrieval device 2 which performs retrieval processing of the technical information based on a retrieval sentence.

The administrative station 3 which registers the technical information into the retrieval device 2.

[0018]Each terminal unit 1 comprises comparatively small computers, such as a personal computer (the following, PC), and the retrieval device 2 comprises comparatively large-sized computers, such as a server.

[0019]Have the retrieval device 2 and the display 2a and printer 2b to the retrieval device 2. The technical information data base (DB) device 4 with which the technical information used as a retrieval object was registered, the pertinent information database device 5 with which the pertinent information on a paper etc.

was registered, and the related term phrase database device 6 with which the related term phrase was registered are connected.

[0020]Each terminal unit 1 and the retrieval device 2 transmit and receive data via communications networks, such as a Local Area Network (following, LAN), the Internet, or a communication line.

[0021]A user inputs a retrieval sentence from the keyboard of the terminal unit 1 for users, etc. Neither number-of-letters restrictions nor character type restrictions are in a retrieval sentence, and arbitrary texts can be inputted into it in each one of language. Even if the formal name of a retrieval object is unknown, it can search by inputting the text explaining a concept or structure. For example, it can refer to inputting the text "the structure and art which are generated using hydrogen and oxygen" as a retrieval sentence. In this case, the technical information about a "fuel cell" is extracted so that it may mention later.

[0022]The retrieval device 2 performs retrieval processing based on the retrieval sentence inputted with the terminal unit 1 for users, and the search results are sent to the terminal unit 1 for users. The administrative station 3 registers the technical information into the retrieval device 2, and the retrieval device 2 transmits management information to the administrative station 3.

[0023]Drawing 2 is a flow chart explaining the processing operation of a 1st embodiment of the technical-information search system of drawing 1. The retrieval sentence is sent to the retrieval device 2 as a user inputs a retrieval sentence from the terminal unit 1 for users (Step S1). The retrieval device 2 extracts the related term contained in a retrieval sentence (Step S2).

[0024]Then, the retrieval device 2 conducts a morphological analysis (Step S3). A morphological analysis extracts the word contained in the text (natural sentence), and specifically extracts a complex NP and a numeral. For example, in the case of a retrieval sentence called "the structure generated using hydrogen and oxygen and art" which were mentioned above, six words and the similar words of those of "hydrogen", "oxygen", "power generation", "structure", "art", and "using" are extracted.

[0025]Then, the dignity (weight) which considered the contents of the technical information for each (search term group) word extracted by the morphological analysis is given, and coded data (vector) changes (step S4). This processing is called indexing of the search term group, and, below, calls a word vector the coded data generated by this processing. For example, it considers that the term which judges that a term, the prefix, the suffix, etc. which occur frequently to any file in a database device have low importance, makes weight low, and occurs frequently only to a specific technical-information file is a word representing the technical information, and it makes weight high.

[0026]After processing of step S4 is completed next, vector processing is performed between the technical information registered into the technical information data base device 4, and the index of a search term group, and similarity is detected (Step S5).

[0027]Here, the word vector which carried out vectorial representation of the technical information registered into the technical information data base device 4 is compared with the word vector which carried out vectorial representation of the search term group of a retrieval sentence, and direction of a vector extracts near information. The nearness of direction of a vector can be expressed with an inner product, and the value of this inner product is expressed with the index of similarity at this embodiment.

[0028]For example, all (weight) of the semantic dignity of six words contained in the retrieval sentence mentioned above, "hydrogen", "oxygen", "power generation", "structure", "art", and "it using", in being equivalent, The document in which the combination of the word nearest to the combination of these words is

included is searched out of the technical information data base device 4.

[0029]Supposing it is the variable which word each became independent of, in the case of the above-mentioned retrieval sentence, it is expressed with a specific 6-dimensional synthesized vector among a number of dimensions. For this reason, an inner product searches the technical information near 1 with Step S5 to direction of this synthesized vector most closely.

[0030]For example, a device which takes out the chemical energy produced as explanation of a "fuel cell" when "substance reacts into the technical information data base device 4 as electrical energy. Usually, although the chemical reaction of hydrogen and oxygen is used, hydrogen is obtained by reforming methanol etc. Supposing there is the technical information ", the word "power generation" is associated from a word called the "electrical and electric equipment" and the "energy" in this technical information. The technical information about the above-mentioned "fuel cell" is extracted from the word "hydrogen" and "oxygen" being in a retrieval sentence by high similarity.

[0031]By the way, Step S2 in the flow chart of drawing 2 for a related term extraction means. In Step S3, Step S54 of drawing 10 for the 1st morphological-analysis means for the 2nd morphological-analysis means. Step S5 of drawing 2 -- an inner-product-computing means -- Step S6 corresponds to a similarity judging means, step S9 corresponds to an output object selecting means, and Steps S10-S12 correspond to a search-results output means, respectively.

[0032]Drawing 3 is a figure showing the retrieving procedure of the technical information data base device 4 to the retrieval sentence mentioned above. Each technical information which should be registered into the technical information data base device 4 is registered into the technical information data base device 4, after a morphological analysis is conducted and a word vector is generated for every word (Steps S21-S24).

[0033]On the other hand, the morphological analysis also of the retrieval sentence which the user inputted is conducted, and a word vector is generated for every word (Steps S25-S27).

[0034]Then, an inner product is calculated between the word vector of a retrieval sentence, and each word vector registered into the technical information data base device 4 (Step S28), and similarity is judged based on the result of an operation (Step S29).

[0035]Much technical information is registered into the technical information data base device 4, an inner product is calculated between search term groups for every technical information, and similarity is detected.

[0036]Thus, in this embodiment, since similarity can be judged by the size of an inner product, similarity can be judged with simply and sufficient accuracy.

[0037]After processing of Step S5 of drawing 2 is completed, the similarity mathematically computed at Step S5 judges whether it is smaller than the standard similarity defined beforehand (Step S6). If there is nothing that exceeds standard similarity in the computed similarity, it will judge that there is no technical information which has relation in a retrieval sentence, and that will be displayed on the terminal unit 1 for users (Step S7).

[0038]On the other hand, similarity extracts the technical information from the technical information data base device 4 about the technical information judged to be more than standard similarity (Step S8). Then, it is chosen in what kind of form search results are displayed on the screen of the terminal unit 1 for users (step S9). For example, when it is chosen that a user displays according to the technical information, processing displayed on the terminal unit 1 for users sequentially from the high technical information of similarity is performed (Step S10).

[0039]Drawing 4 is a figure showing the example which displays search results according to the technical information. In a screen, the field W1 which displays the retrieval sentence which the user inputted, the field W2 which displays a technical-information extraction result list, and field W3 which displays the detailed content of the technical information are provided.

[0040]The name of the technical information is displayed in a list on a technical-information extraction result list by the high order of similarity. Similarity is numerically displayed so that it can be grasped whether each technical information is how much similar to the retrieval sentence at a glance. It is an index which shows that they are the contents near a retrieval sentence, and shows that it is useful information for a retrieving person, so that similarity is high.

[0041]Instead of displaying similarity numerically, as shown in drawing 5, it may display with a bar graph etc. If it displays in a graph, the difference in the similarity of each technical information can be grasped easily.

[0042]If a user chooses the arbitrary technical information in a technical-information extraction result list with a mouse etc., the detailed content corresponding to the technical information will be displayed on right-hand side viewing-area W3. Not only the bibliographic information on the technical information but the whole sentence information and pertinent information which the user chose may be displayed on this field.

[0043]setting to Step S6 of drawing 2 on the other hand -- a user -- a pertinent art classification (for example, International Patent Classification IPC) -- when displaying independently is chosen, processing displayed on the terminal unit 1 for users sequentially from the high pertinent art classification of similarity is performed (Step S8).

[0044]Drawing 6 is a detail flowchart of processing of Step S8 of drawing 2. First, similarity detection according to technical information is performed like Step S7 (Step S31). Then, the pertinent art classification of each technical information by which similarity detection was carried out is detected (Step S32).

[0045]Then, the average value and dispersion of similarity are calculated according to a pertinent art classification (Step S33). Then, the list which arranged the pertinent art classification in the high order of similarity is created (Step S34).

[0046]Drawing 7 is a figure showing an example of the search-results display screen corresponding to Step S8 of drawing 2. The difference from drawing 4 is the point of displaying search results according to a pertinent art classification. Search results are totaled and displayed according to a pertinent art classification, and the user can grasp the technical classification which has relation in a retrieval sentence, can specify a technical classification, and can perform further search.

[0047]On the other hand, in Step S6 of drawing 2, when it is chosen that a user displays according to a related company, processing displayed on the terminal unit 1 for users sequentially from the high related company of average similarity is performed (step S9).

[0048]Drawing 8 is a detail flowchart of processing of step S9. After performing similarity detection according to technical information (Step S41), the related company of each technical information by which similarity detection was carried out is detected (Step S42).

[0049]Then, the average value and dispersion of similarity are calculated according to a related company (Step S43). Then, the list which arranged the related company in the high order of similarity is created (Step S44).

[0050]Drawing 9 is a figure showing an example of the search-results display screen corresponding to step S9 of drawing 2. The difference from drawing 4 is the point of totaling and displaying search results

according to a related company. By displaying search results according to a related company, the user can grasp the company name which has relation in a retrieval sentence at a glance.

[0051]Since utility value increases so that there is much technical information registered into the technical information data base device 4, it enables it to register the new technical information into the technical information data base device 4 one by one in this embodiment.

[0052]Drawing 10 is a flow chart which shows the procedure of the technical-information registration processing which the administrative station 3 performs. First, the technical information which should be registered is incorporated (Step S51). Then, the data format of the incorporated technical information is transformed into the data format corresponding to the technical information data base device 4 (Step S52). Here, tag information is mainly changed.

[0053]Drawing 11 is a figure showing a list of the tag which each technical information registered into the technical information data base device 4 has, and shows the example of the tag for patent documents.

[0054]<doc> which shows the head of a document to the tag for patent documents, <title\> which shows the title of a document, There are <createddate\> which shows an application year, <au\> which shows a patentee, <keywordt\> which shows a keyword, <body> which shows the text, </doc> which shows the last of a document, etc.

[0055]Drawing 12 is a figure showing an example of the technical information registered into the technical information data base device 4, and shows the example of the patent document. Like a graphic display, since each technical information is registered into the technical information data base device 4 where a tag is attached, it can spy out desired engineering documentation simply and promptly by searching based on these tags.

[0056]An end of the tag conversion process of Step S52 of drawing 10 will create the technical-information data for search next (Step S53). Then, after conducting a morphological analysis by the same technique as Step S3 of drawing 2 and extracting a related term (Step S54), weighting which seasoned each related term with importance etc. is performed (Step S55). At this time, the digital data which expressed each related term in vector is generated, and this data is registered into the technical information data base device 4 (Step S56).

[0057]Drawing 13 - drawing 15 are flow charts which show the procedure of the tag conversion process performed at Step S52 of drawing 10. First, the error log file which records error information is opened (Step S71). Next, the processing log file which records processing information, including processing start time etc., is opened (Step S72).

[0058]Next, processing start time is recorded on a processing log (Step S73). Next, the patent document data in CD-ROM on which the patent document was recorded is read one by one (Step S74). Next, various tag variables are initialized (Step S75). variable $author which specifically expresses an author name -- a null -- <createddate> is inputted into variable $c_date which expresses a creation date for <title> to variable $title which expresses a title for <keywordt> to variable $kt which expresses a keyword for data.

[0059]Next, after initializing a flag (Step S76), a patent data file is opened (Step S77), and patent document data is read a party every (Step S78). It is judged whether the term of "the name of an invention" enclosed in the glue-block parenthesis exists in a read-out Ta line (Step S79). If this term exists, the concrete name of the name of an invention will be substituted for variable $title (Step S80).

[0060]When judged with NO at Step S79, or when processing of Step S80 is completed, it is judged whether

the term of the "filing date" enclosed in the glue-block parenthesis exists in a read-out Ta line (Step S81). If this term exists, after converting the full-width character showing the date with the half-width character of the form of yyyy/mm/dd (Step S82), this half-width character is substituted for variable $c_date (Step S83).

[0061]When judged with NO at Step S81, or when processing of Step S83 is completed, it is judged whether the term of the "patentee" or the "applicant" enclosed in the glue-block parenthesis exists in a read-out Ta line (Step S84 of drawing 14). If this term exists, after skipping to the line in which the term of "the name or name" enclosed in the glue-block parenthesis exists, the concrete name or name indicated in that line is substituted for variable $author (Step S85).

[0062]When judged with NO at Step S84, or when processing of Step S85 is completed, One [ keyword Field Flagg ] (Step S87) if it judges whether the term of "International Patent Classification" enclosed in the glue-block parenthesis exists (Step S86) and this term exists.

[0063]When judged with NO at Step S86, or when Step S87 is completed, it is judged whether tag conversion was performed to the final line of a patent data file (Step S88). When tag conversion is omitted to the final line yet, the processing after Step S78 is repeated.

[0064]When tag conversion is performed to a final line, it is judged whether a tag omission exists (Step S89). An error log is outputted when a tag omission exists (Step S90).

[0065]When judged with NO at Step S89, or when processing of Step S90 is completed, it is judged whether an output directory exists (Step S91). When an output directory does not exist, a directory is created the first stage (Step S92).

[0066]When judged with YES at Step S91, or when processing of Step S92 is completed, an object year file is opened (Step S93). First, a start tag <doc> is opened (Step S94). Next, the extracted tag is outputted to an object year file (Step S95). Here, the tags to extract are <title>, <createddate>, <au>, and <keuwordt>, for example.

[0067]Next, the whole contents of patent document data are outputted after the tag <body> of the main part of a patent document (Step S96 of drawing 15). Next, the file is closed after outputting an end tag </doc> to an object year file (Step S97).

[0068]Next, it is judged whether tag conversion and the output to the file by year were completed about all the patent document data recorded on CD-ROM (Step S98). Next, the name and number of CD-ROM which the registration to tag conversion and the file by year ended are recorded on a processing log (Step S99).

[0069]Next, it is judged whether the processing about all the CD-ROMs was completed (Step S100). If there is a CD-ROM which processing has not ended yet, processing after Step S74 will be performed (Step S101). Next, a processing log file and an error log file are closed (Step S102).

[0070]Drawing 16 and drawing 17 are flow charts which show the detailed content of the retrieval processing explained with the flow chart of drawing 2. First, after opening the input file on which the search string is recorded (Step S121), the output file which records search results is opened (Step S122).

[0071]Next, a search string is read after setting the number of reference documents which stored the technical information which is a retrieval object (Step S123) (Step S124). Next, it is judged whether the read search string has an error (Step S125). In being errorless, it generates the vector of a search string (Step S126). Next, retrieval processing of the technical information is performed based on the generated vector (Step S127).

[0072]Next, it is judged whether the degree of association which the user set up is zero (Step S128). If a

degree of association is larger than zero, related term extraction will be performed based on the degree of association (Step S129). Next, the technical information is searched based on the extracted related term (Step S130).

[0073]When judged with YES at Step S128, when processing of Step S130 is completed, After performing a total and alignment of the IPC code (Step S131), A total and alignment of a patentee are performed (Step S132), and a total and alignment of an application year are performed after that (Step S133), next a total and alignment result of Steps S130-S132 are outputted (Step S134).

[0074]Next, it is judged whether the numerical value of the summary rate which the user set up is zero (Step S135). Here, a summary rate is a ratio of the number of summary sentence chapters to the number of whole sentence chapters. Abstract processing is performed when the numerical value of a summary rate is larger than zero (Step S136).

[0075]Next, it judges whether all the search results were obtained (Step S137), and if there are some which have not been obtained yet, the processing after Step S124 will be repeated. When all the search results are obtained, processing is ended after outputting the result (Step S138).

[0076]Thus, in a 1st embodiment, conduct the morphological analysis of the retrieval sentence which the user inputted, and a word vector is generated, Since similarity is judged based on the result of having calculated the inner product with the word vector of each technical information registered into the technical information data base device 4, similarity with each technical information can be judged with simply and sufficient accuracy.

[0077]In order to display search results with a gestalt which the height of similarity understands, even if two or more technical information is searched, the height of the similarity of each technical information can be grasped objective.

[0078]In a 1st embodiment, since high-precision search can be performed even if the technical keyword is not contained in the retrieval sentence, a retrieval sentence can be created in the same form as the usual text, and since **** becomes there is not in creation of a retrieval sentence less about skill, a user's user-friendliness improves.

[0079](A 2nd embodiment) When a 2nd embodiment displays search results, it displays advertisement information with relation on a retrieval sentence collectively.

[0080]Drawing 18 is a block diagram showing the outline composition of a 2nd embodiment of the technical-information search system concerning this invention. In drawing 18, identical codes are given to the component part which is common in drawing 1, and, below, it explains focusing on a point of difference.

[0081]The technical-information search system of drawing 18 is provided with the terminal unit 7 for advertisers which provides a product (goods) and the advertisement information of service. This terminal unit 7 for advertisers comprises a PC etc., and performs transmission and reception of the retrieval device 2 and data via communications networks, such as the Internet.

[0082]The retrieval device 2 of drawing 18 has the advertisement information database device (advertisement information DB) 9 which registers the advertisement information provided from the terminal unit 7 for advertisers. This database device 9 may be stored in the same recorder as the technical information data base device 4 and the pertinent information database device 5 which were mentioned above, and may form a recorder separately, respectively.

[0083]Drawing 19 is a flow chart which shows the outline composition of a 2nd embodiment of the technical-

information search system of drawing 18. The procedure of registering the new technical information into the technical information data base device 4, the retrieving procedure, and the procedure of registering new advertisement information into the advertisement information database device 9 are shown in drawing 19. Since the procedure (Steps S151-S156) registered into the technical information data base device 4 is [ among these ] the same as drawing 10, explanation is omitted.

[0084]If an advertiser inputs a product and the advertisement information of service via a sponsor terminal device (Step S157), the retrieval device 2 will conduct the morphological analysis of the input (Step S158), will consider importance etc., and will perform weighting processing (indexing) of advertisement information (Step S159). The word vector about advertisement information is generated and, more specifically, it registers with the advertisement information database device 9 (Step S160).

[0085]An advertiser transmits the document which recorded the required matter according to the format shown in drawing 15 to a retrieval device via a communications network, when registering new advertisement information.

[0086]On the other hand, if a user inputs a retrieval sentence via the terminal unit 1 for users (Step S161), the retrieval device 2 will conduct the morphological analysis of the retrieval sentence (Step S162, S163), will consider importance etc., and will perform weighting processing of a search term group (Step S164).

[0087]Then, while the retrieval device 2 calculates an inner product between the word vector about a search term group, and each word vector registered into the technical information data base device 4, An inner product is calculated between the word vector about a search term, and each word vector registered into the advertisement information database device 9, and similarity is detected (Step S165).

[0088]Then, the retrieval device 2 judges similarity based on the result of an operation of an inner product (Step S166). It is judged whether the result of an operation of an inner product is more specifically smaller than the standard similarity inputted beforehand. When all the results of an operation of an inner product are smaller than standard similarity, it indicates that judges that similarity is low and there is no applicable technical information (Step S167).

[0089]On the other hand, if there is a thing more than standard similarity among the results of an operation of an inner product, the advertisement information which extracts the applicable technical information from the technical information data base device 4, and has relation in the technical information will be extracted from the advertisement information database device 9 (Step S168).

[0090]Then, after making a user choose the display style of search results (Step S169), the technical information and the advertisement information which has relation in the technical information are displayed a technical-information exception, a pertinent art classification exception, or according to a related company (Step S170, S171, S172).

[0091]By the way, in Step S162 of drawing 19, Step S163 for a related term extraction means for the 1st morphological-analysis means. In Step S154, Step S158 for the 2nd morphological-analysis means for the 3rd morphological-analysis means. Step S156 -- the 1st information accumulation means -- Step S160 -- the 2nd information accumulation means -- Step S165 corresponds to an inner-product-computing means, Step S166 corresponds to a similarity judging means, and Steps S170-S172 correspond to a search-results output means, respectively.

[0092]Drawing 20 is a figure showing the example of a screen display of the search results displayed on the terminal unit 1 for users. Like a graphic display, the viewing area of advertisement information is provided in

addition to the viewing area of a retrieval sentence, the viewing area of the extraction result list of technical information, and the viewing area of the detailed content of the technical information. The size of this viewing area may be arbitrarily changed according to the number of advertisement information.

[0093]If advertising display size is changed according to the height of similarity when displaying advertisement information, the higher advertisement information of a user's concern can display greatly, and an advertising effect can be heightened further. Since the product with which he is interested also for a user, or the advertisement information of service is displayed greatly, a retrieval object can be easily attained by checking this advertisement information.

[0094]When a user chooses with a mouse etc. the specific advertisement information displayed in the screen of drawing 20, it is desirable to display the product about the advertisement information or the outline of service, or to display homepages, such as a company which provides the advertisement information. Thereby, only when a user needs, the detailed content of advertisement information can be displayed.

[0095]By the way, the quantitative effect by advertising printing is expressed with the following (1) types.

[0096]a=(bxc)/sigmabxd -- (1)

The similarity at the time of the advertisement concerned being displayed here as for the number of times and c as which the advertisement concerned was displayed [ a ] on the screen as for an advertising effectiveness index and b, and d are the number of times by which the banner of the advertisement concerned was chosen with the mouse etc. and the outline of the advertisement concerned was referred to.

[0097](1) It is shown that advertising effectiveness is so large that the value of a computed by a formula is large. Therefore, if an advertising printing fee is set up based on the value of this a, the claim according to advertising effectiveness can be performed to each advertiser. The above-mentioned (1) type is equivalent to a printing setting means.

[0098]Also about the objective market value of the technical information registered into the technical information data base device 4, it can measure quantitatively in a similar manner, and the quantitative effect in this case is expressed with the following (2) types.

[0099]S=(txu)/sigmatxv -- (2)

The similarity at the time of the market value index of the extraction technical information, the number of times by which the technical information concerned was displayed in a list on the screen as for t, and the technical information concerned being displayed here as for S, and v are the number of times by which the listed technical information concerned was chosen with the mouse etc., and the outline of the technical information concerned was referred to.

[0100]This S can be used as a standard of value in case license dealings are actually conducted. Since it is thought that it is worthy so that the value of this S is large, about the large technical information of S, it is desirable to carry out highlighting of the title information and to call a user's attention. The above-mentioned (2) types are equivalent to a market value estimation means.

[0101]The advertiser can propose to advertising printing on on-line via the Internet etc. Drawing 21 is a figure showing an example of the application screen of advertising printing on on-line. Like a graphic display, an advertiser proposes to advertising printing via the Internet etc., after inputting a catch copy, a product name, a company name, a contact, a product summary, etc. into each column. The retrieval device 2 stores the advertisement information from an advertiser in the advertisement information database device 9.

[0102]It not only displaying the technical information which has relation in the retrieval sentence which the

user inputted in a 2nd embodiment in order of similarity, but having displayed the advertisement information which has relation in the technical information Thus, a sake, Only the user who is interested in a specific product and service can be provided with an advertisement, and an advertising effect can be heightened. That is, only the user considered to be interested in advertisement information in order to provide a user with the advertisement information which has relation in search results can be provided with advertisement information, and the goods by an advertisement and the sale of service can be improved rather than advertising to many and unspecified persons.

[0103]In order to provide a user with advertisement information in the stage where search results were obtained, it is not necessary to provide a user with advertisement information vainly, and stops making a user trouble.

[0104]Since it can propose to advertisement information on-line, while being able to give facilities to an advertiser, a change of advertisement information can also be made easily.

[0105]When performing new registration of advertisement information, he is trying to display on the terminal unit 7 for advertisers the product information index shown in drawing 21 in this embodiment. The product information index of drawing 21 has each column of a catch copy, a product or a service name, a company name, a contact, and a product summary. The information inputted into the product information index is sent to the retrieval device 2, and is registered into the advertisement information database device 9.

[0106]

[Effect of the Invention]As explained to details above, according to this invention, since search results were outputted taking the similarity of a retrieval sentence and the retrieval object technical information into consideration, it can be grasped whether each retrieval object technical information is how much similar to the retrieval sentence easily and correctly. Thereby, quick and the increase in efficiency of the examination business of the examiner who performs a patent examination based on advanced-technology literature can be attained. While being able to attain the precision improvement and the increase in efficiency of patent trend analysis (for example, research-and-development strategy planning, investigation of the status of development of the other company, etc.) of a company, utilization of inter-technology analysis can be attained.

[0107]According to this invention, since the similarity with a retrieval sentence provided high advertisement information automatically, only those who are interested in an advertisement can be provided with an advertisement, and high advertising printing of the opposite cost effect by direct marketing can be realized. Since those who are not interested in an advertisement are no longer provided with advertisement information, it stops making the user of a search service trouble.

[0108]Only the part which includes advertisement information in a search service can set up the fee of a search service cheaply, and can plan promotion of utilization of a search service. Advertising effectiveness can be heightened by changing the contents of advertisement information, etc. according to similarity.

[Translation done.]